Pattern, process and hypothesis testing

Robert H. Gardner

Philip A. Townsend

Appalachian Laboratory

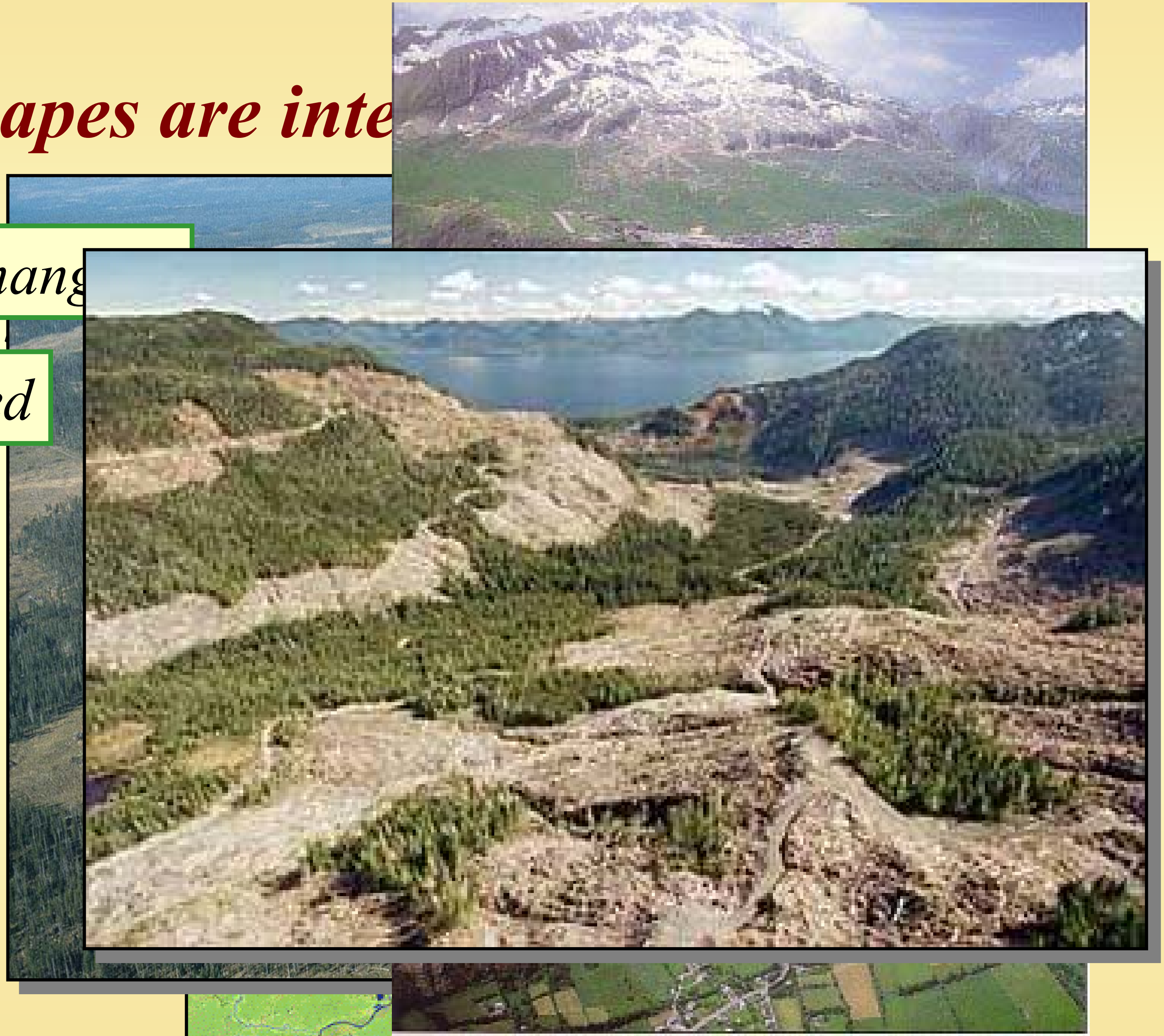
University of Maryland Center for Environmental Science

Landscapes are inte

ever chang

they are

disturbed



Landscape metrics are uninteresting

unusable

$$M_i = \sum_{k=1}^{m} \left(\sum_{j=1}^{n} d_{ij} \right) / n_i$$

$$P_x = \frac{-\sum_{k=1}^{m}\left[\left(\frac{e_{ik}}{\sum_{k=1}^{m} e_{ik}}\right)\ln\left(\frac{e_{ik}}{\sum_{k=1}^{m} e_{ik}}\right)\right]}{\ln (m-1)} 100$$

Why quantify landscape pattern?

To determine if:

- pattern changed over time?
- compare different landscapes
- evaluate consequences of management options, policies, conservation practices
- anticipate changes in process
  - e.g., species loss

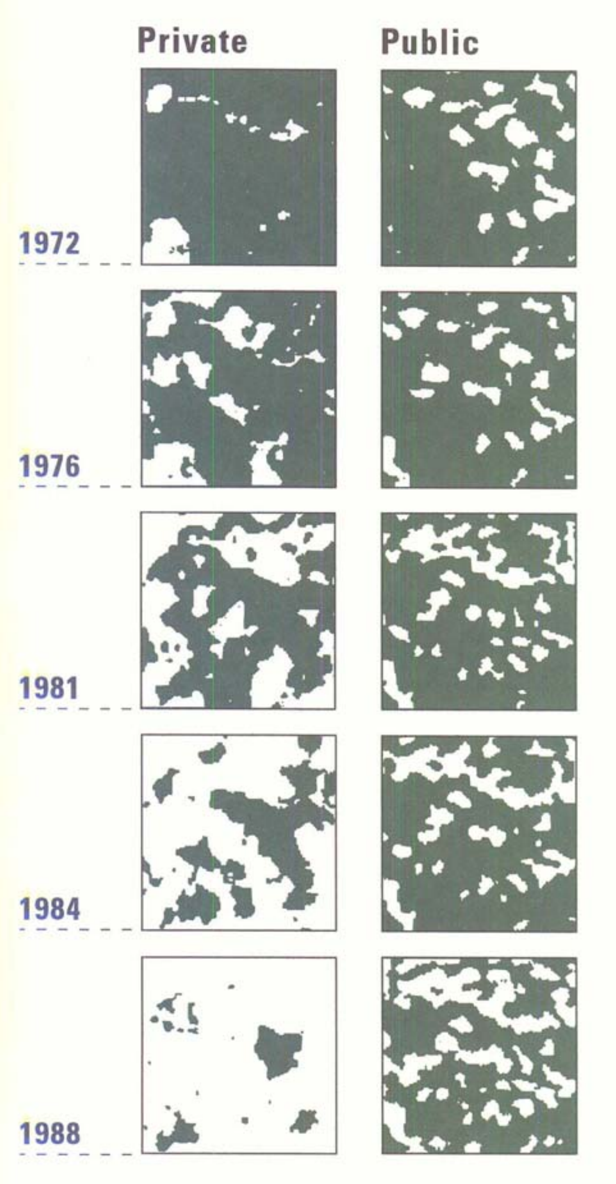


Figure 5.3

For this presentation …

- A brief (personal) history of landscape metrics
- Problems and pitfalls using landscape metrics
- One illustration of pattern-process analysis
- Some recommendations

Spatial analysis in ecology

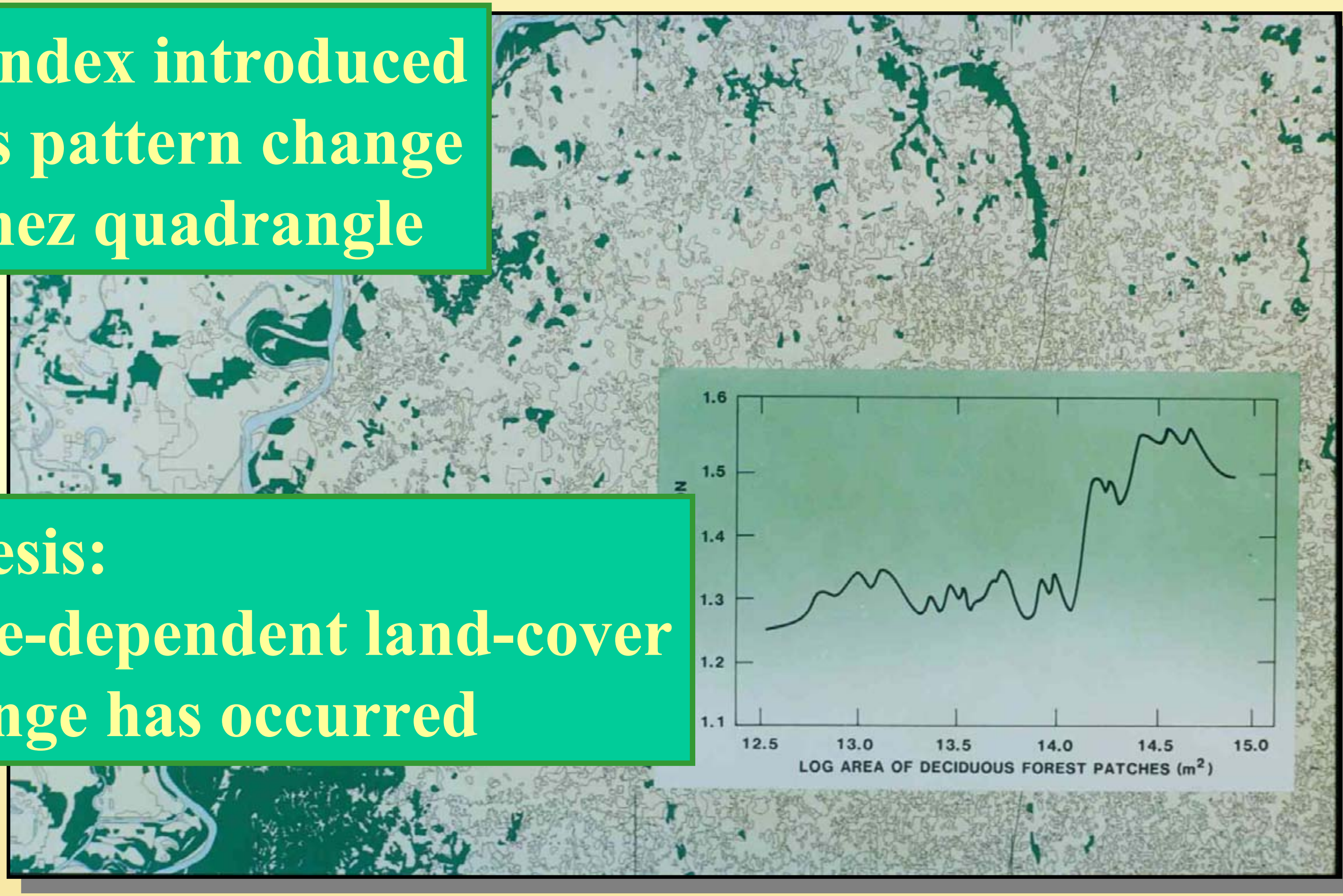
- Spatial systems have been of considerable interest
  - long before computers were available
- Pielou's 1963 book provides a summary, including:
  - *Spatial patterns of discrete quantities*
  - *Measurement of aggregation*
  - *Individuals in a continuum*
  - *Patterns measured by distance sampling*
  - *Two-phase mosaics (maps)*

Map availability

- Kuchler's (1969) ***potential vegetation maps***
  - peaked (U. S.) interest in broad-scale vegetation patterns
  - but left the question: "What is really there?"
- The advent of ***remote imagery*** and ***computer mapping*** provided extensive land-cover maps
  - NASA high-altitude U2 flights (1973) provided continental coverage
  - USGS produced digital data (LUDA, circa 1983) from NASA's aerial coverages
  - BUT "How (and why) do you analyze these patterns?"

Krummel et al (1987) took advantage of digital data and computational advances

fractal index introduced to assess pattern change in Natchez quadrangle



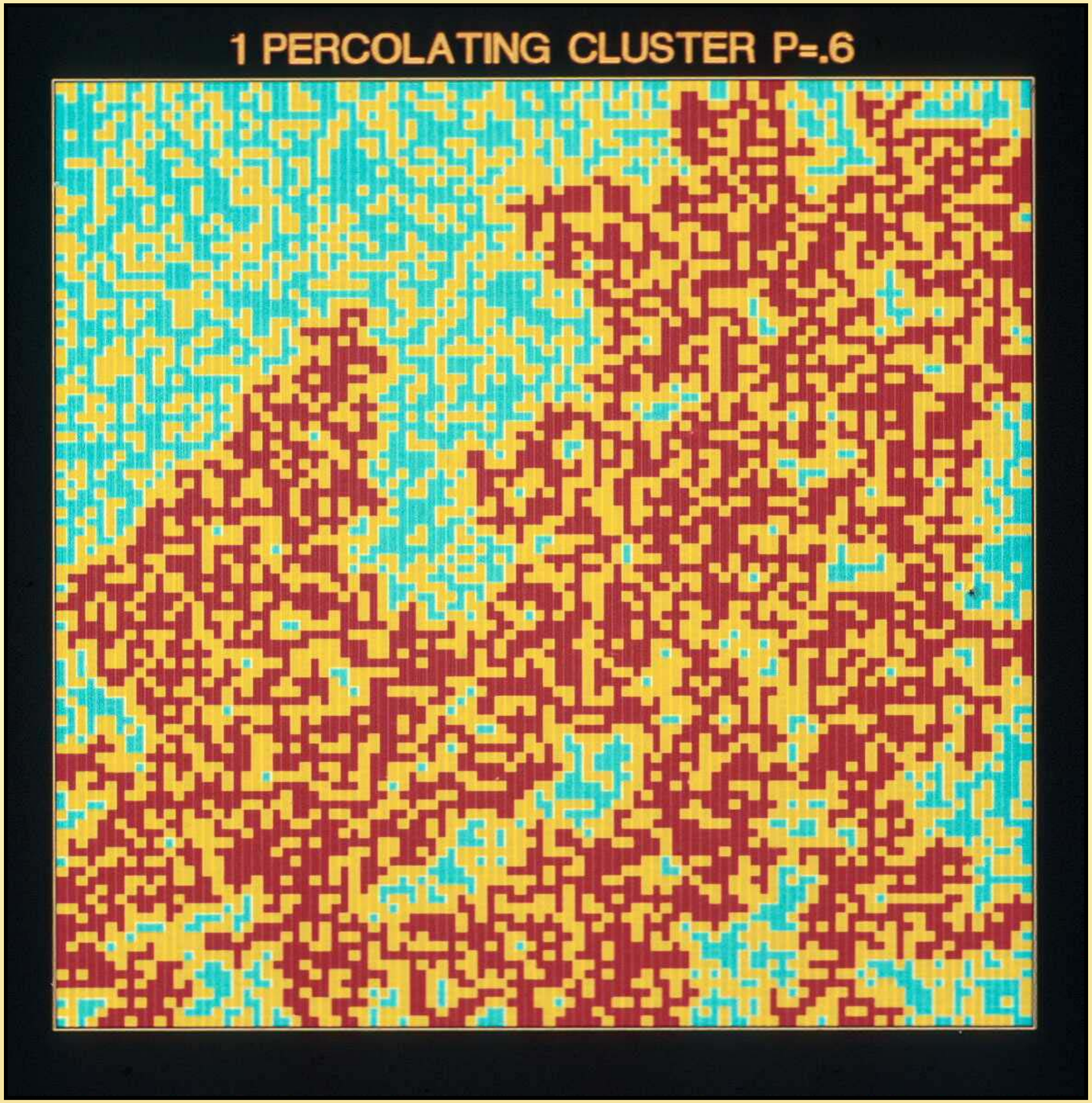
Hypothesis:

scale-dependent land-cover change has occurred

Patterns appeared similar to those generated from percolation theory

Therefore:

percolation-based indices developed



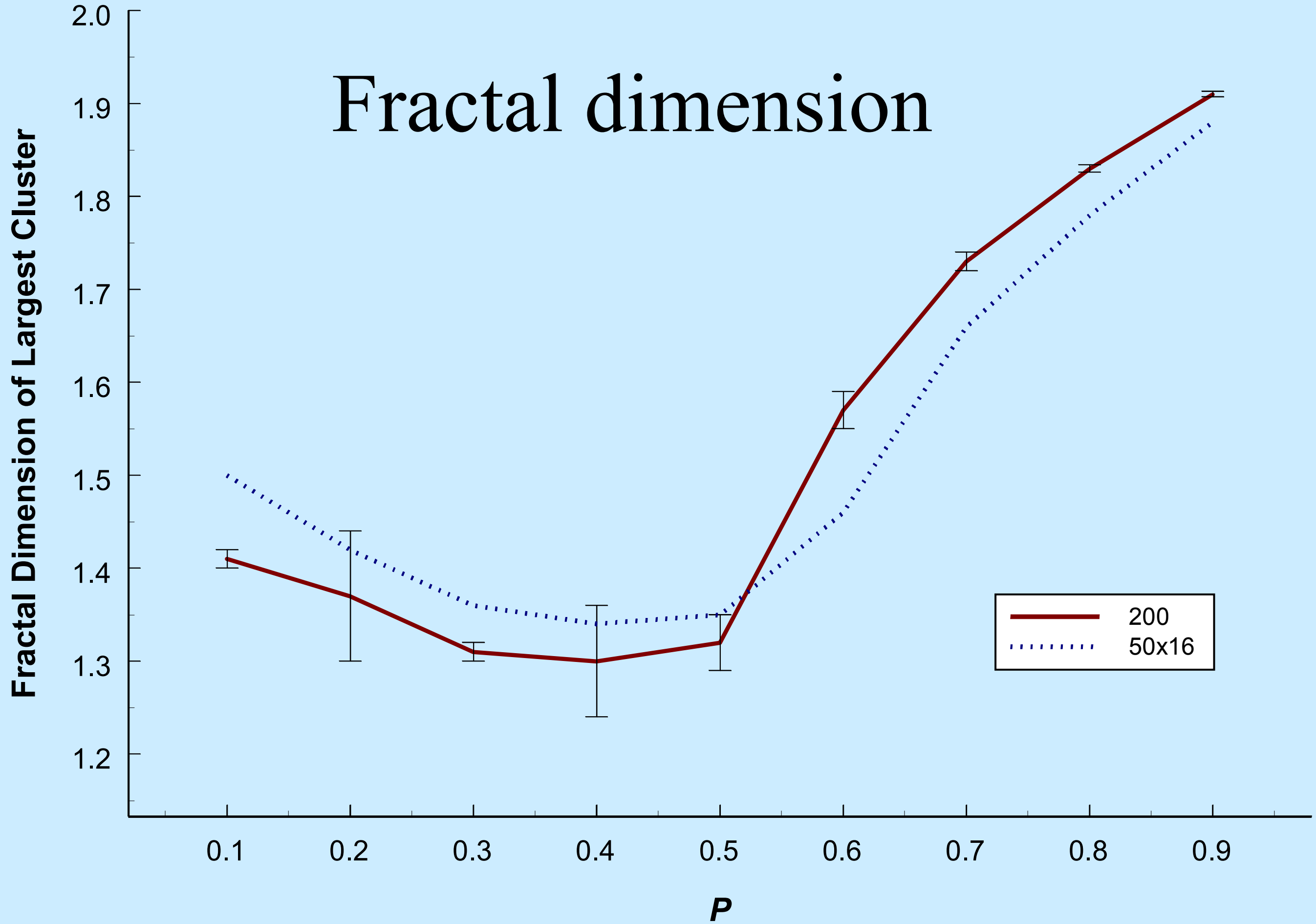
Many metrics employed simply because it was possible

- O'Neill's 1989 paper in Landscape Ecology
  - identified numerous possible metrics
- Several software packages became available to the *landscape cognizati*
  - Monica Turner's early program, RULE available since '91
  - but FRAGSTATS provided first comprehensive calculation of ~100 indices
    - ❖ widely available since ~1993

What we now know

- there are only a few fundamental variables affecting landscape pattern
  – other metrics are redundant
- the number of land-cover classes dramatically affects the analysis results
  – aggregated maps, or those developed with different classification rules, should not be compared
- map scale (grain and extent) must be carefully specified
  – extent affects boundaries

Examining behavior of indices with different sized maps



Example: estimating fractals

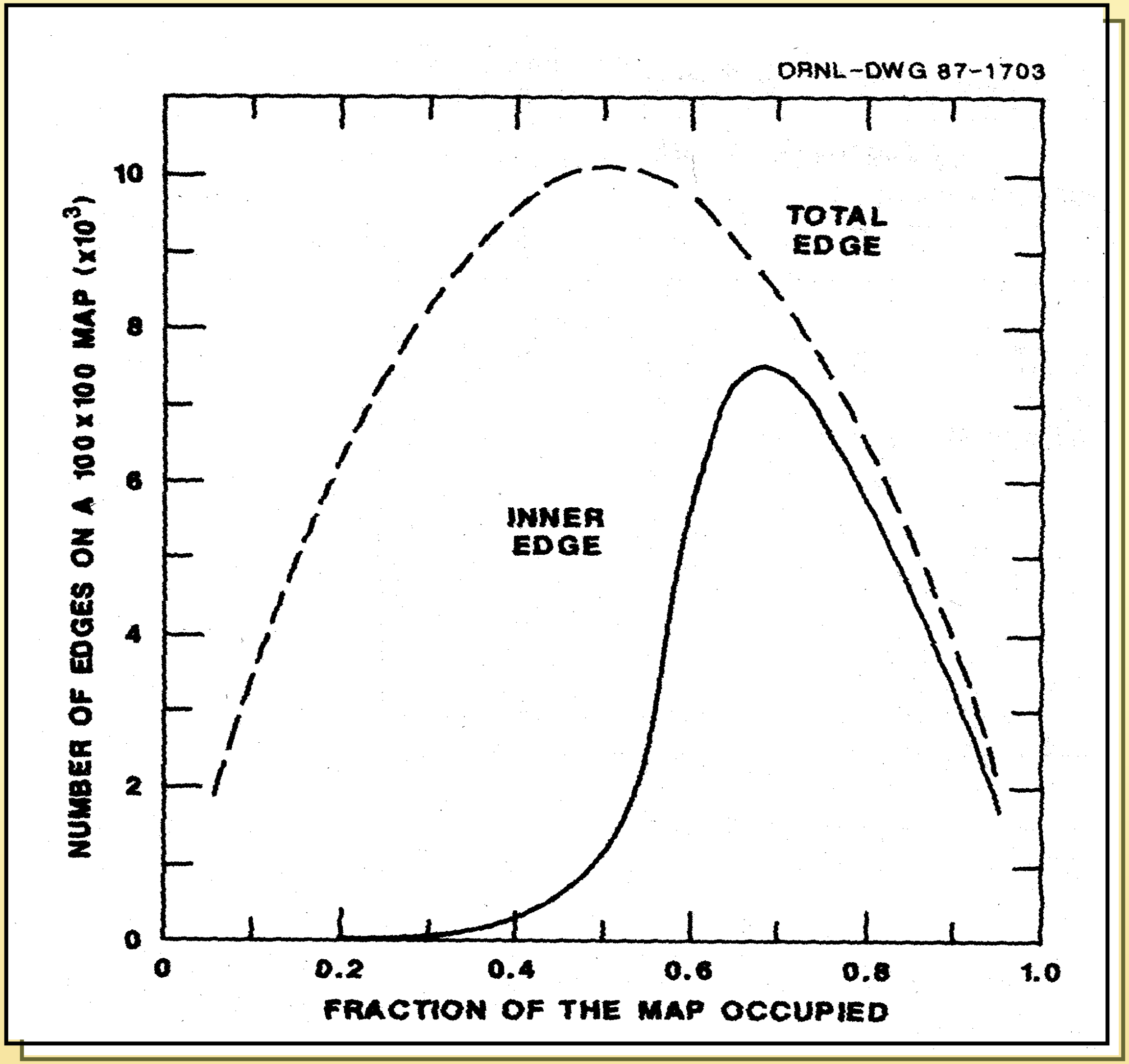
- Log-log regressions are unreliable
  - ***small clusters bias estimates*** because fractals for objects < 100 sites are biased by the shape of the grid
- Mean fractal of patch-wise averages are also unreliable
  - ***because distribution not normal***
- Most reliable estimate
  - the fractal dimension ***of the largest patch***
  - this is what the physicists do and what RULE reports?

Five pitfalls of metric diversity

1. Type I error is high when multiple indices are simultaneously evaluated (***the Fragstats pitfall***)
2. Lack of algebraic rigor in the development and analysis of metrics
3. Failure to demonstrate that metric is sensitive to structural changes ***it is intended to measure***
   - monotonic relationships often not established
   - pattern-process dependencies are rarely demonstrated
4. Error analysis (confidence intervals) seldom estimated
5. Hypothesis testing (how far does result differ from expected) is rarely implemented

Non-monotonic relationships

function must be known in order to extrapolate or compare



Hypothesis testing should be the goal

- Why hypothesis testing is necessary
- Three examples (the null hypothesis)
  - no significant change has occurred (compare self through time)
  - pattern does not deviate from expectation
  - different landscapes have similar structures

Example: Dispersal within corridors

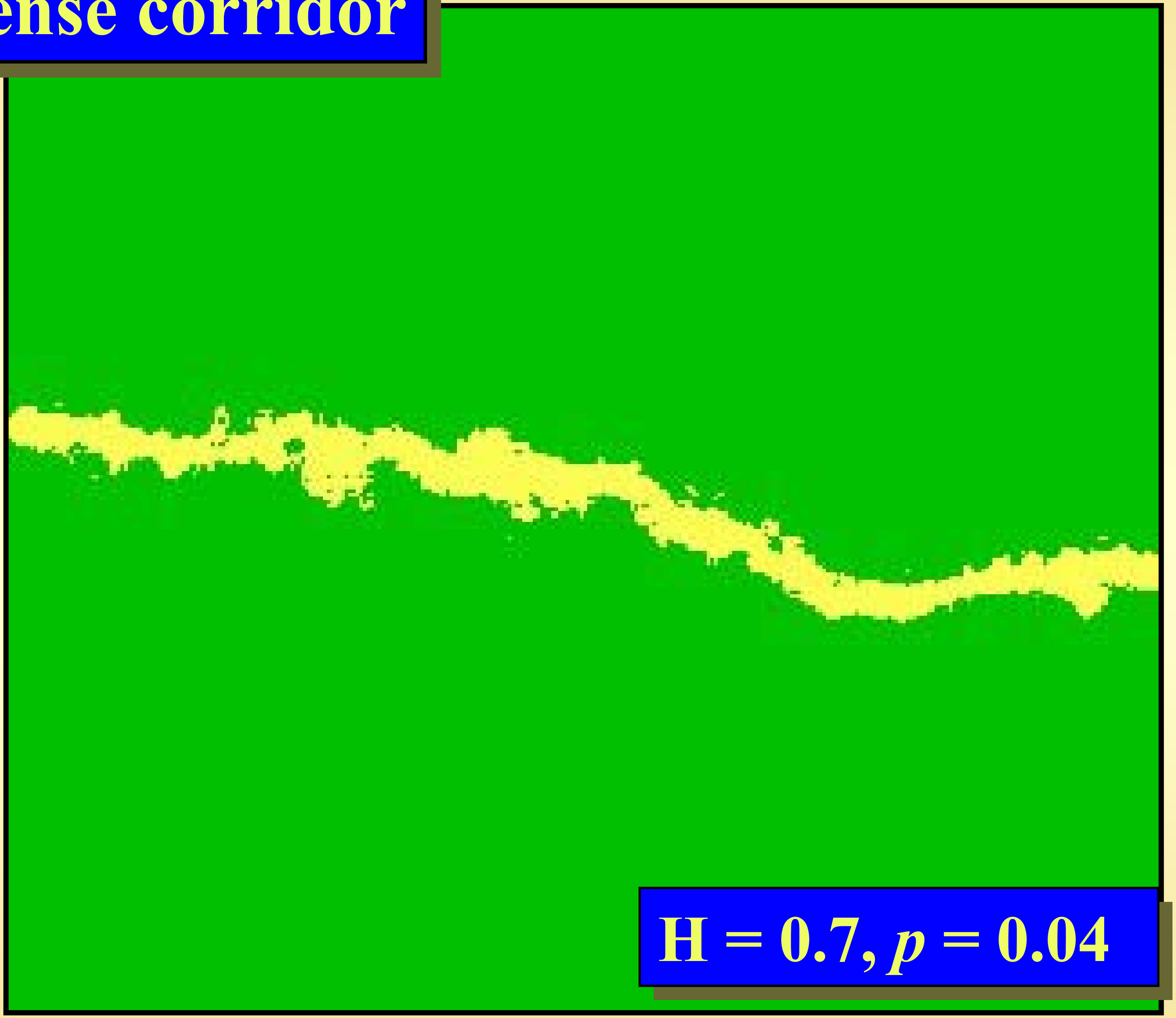
- Dispersal success is affected by landscape pattern
  - as fragmentation increases, dispersal events become more infrequent
- It is poorly described for most species of plants and animals
  - ***but is really important for understanding and predicting persistence, invasions, etc.***
- Will corridors mitigate fragmentation effects?

Simulations using CAPS

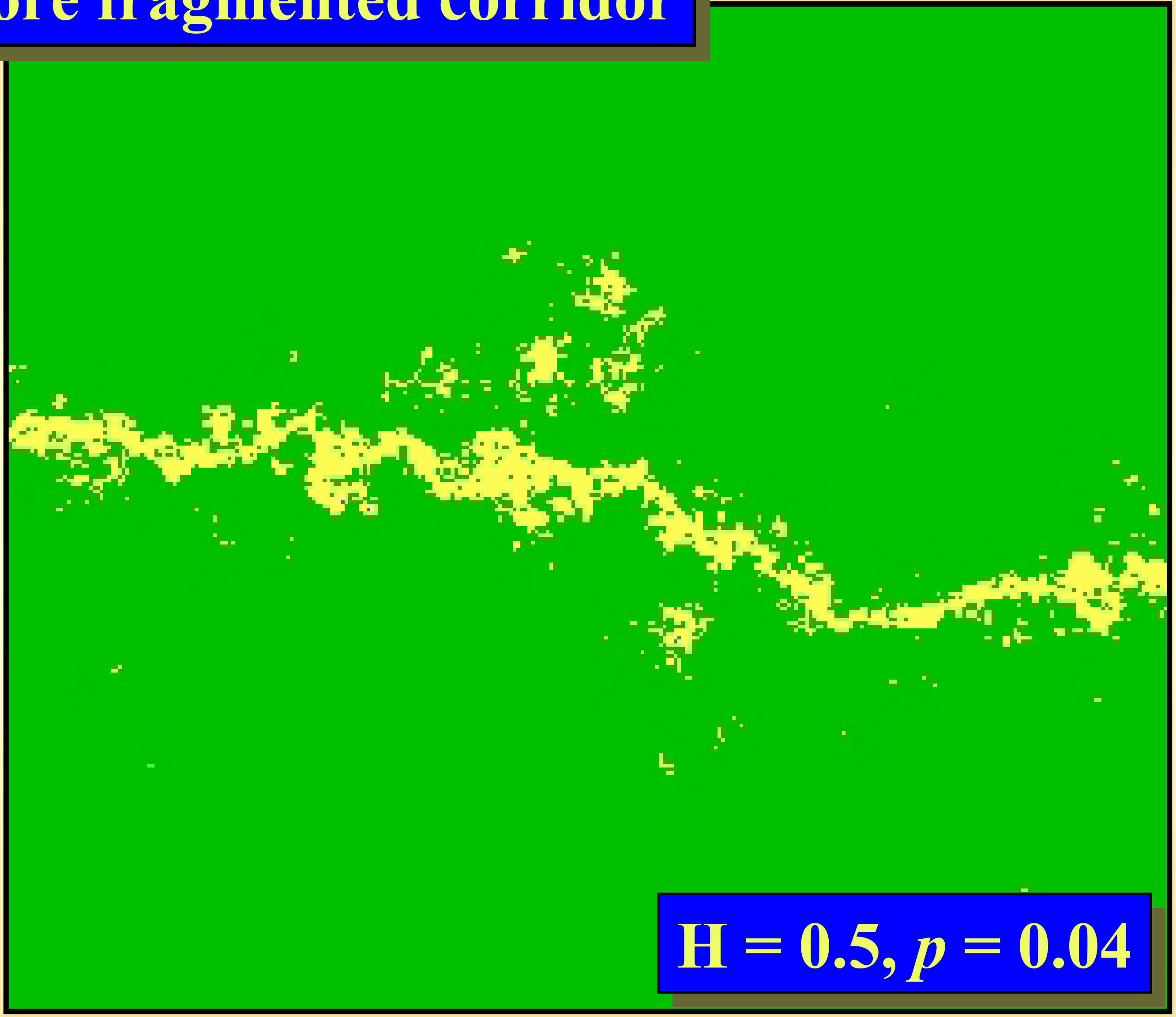
1. Randomly generates corridors
   - via fractal algorithm
     - patterns controlled with H and p
       - H = autocorrelation
       - p = habitat amount
   - allows replicate simulations

(Plotnick and Gardner, 2002)

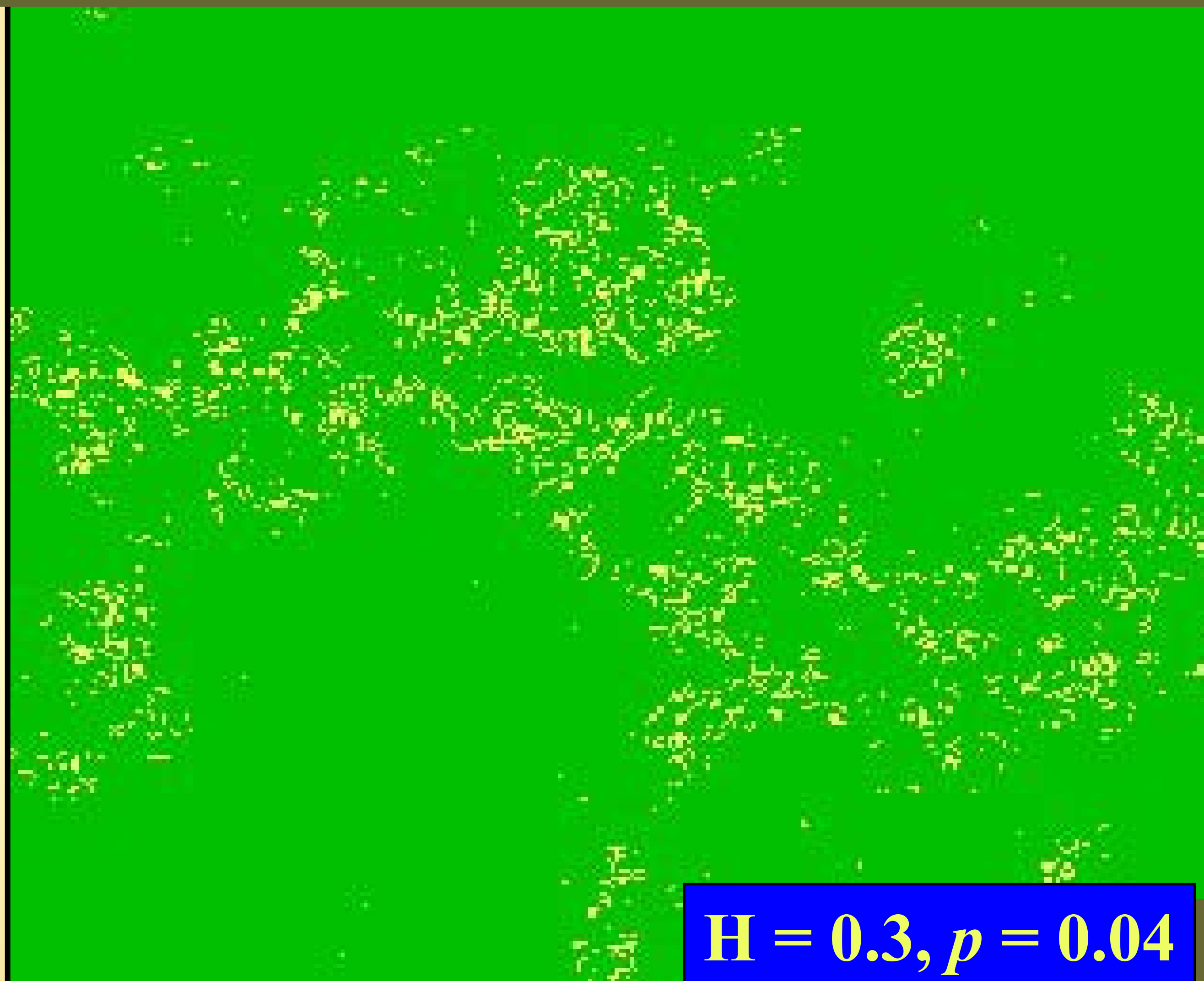
Dense corridor



More fragmented corridor



More fragmented, but same “amount” of habitat



H = 0.3, p = 0.04

How CAPS works

1. Generate map (gridded)
2. Specify species characteristics
   - life span (fixed)
   - **relative fecundity (varied)**
   - dispersal kernel (fixed)
   - **habitat (niche) preferences (varied)**

Competition via seed lottery

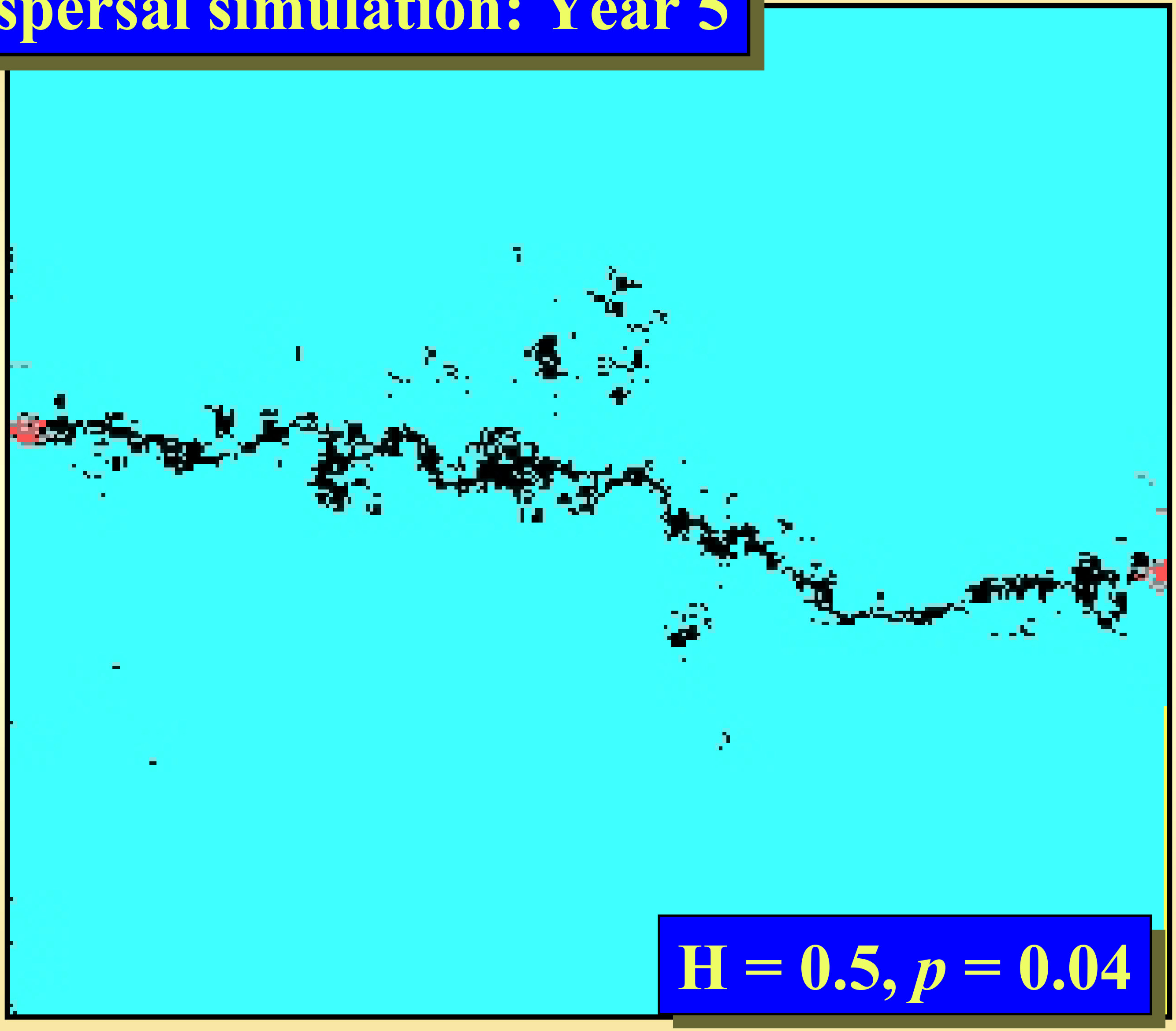
for each map location (i,j)

1. *spread propagules to neighboring cells*
2. *determine if habitat conditions suitable for germination*
3. *randomly select winner based on local seed abundance and competitive ability*

How CAPS works

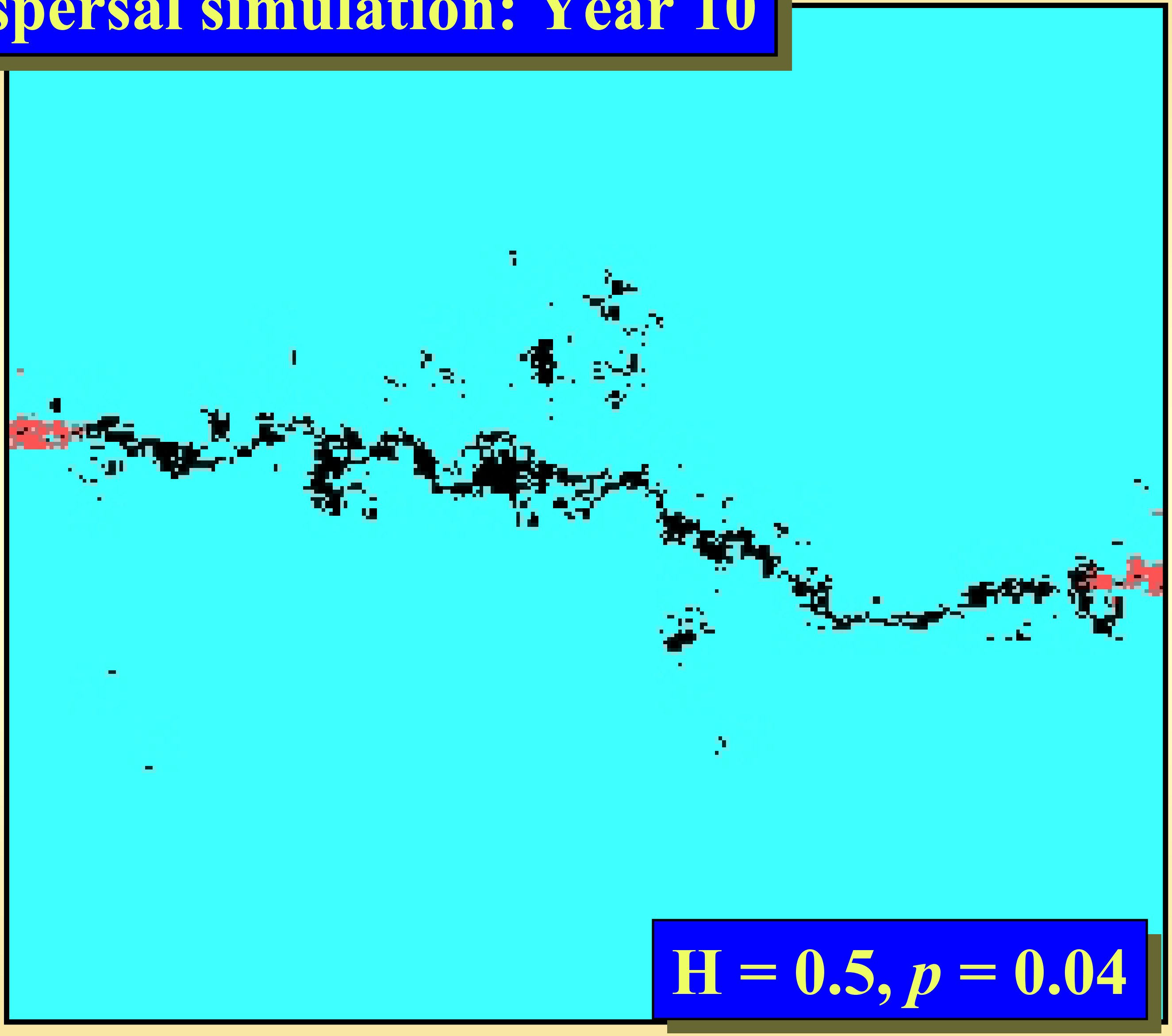
1. Generate map (gridded)
2. Specify species characteristics
3. Initialize map
4. Run simulation
   - Measure rate of movement (***v***)
   - Spatial pattern of occupied habitat (***S***)

Dispersal simulation: Year 5



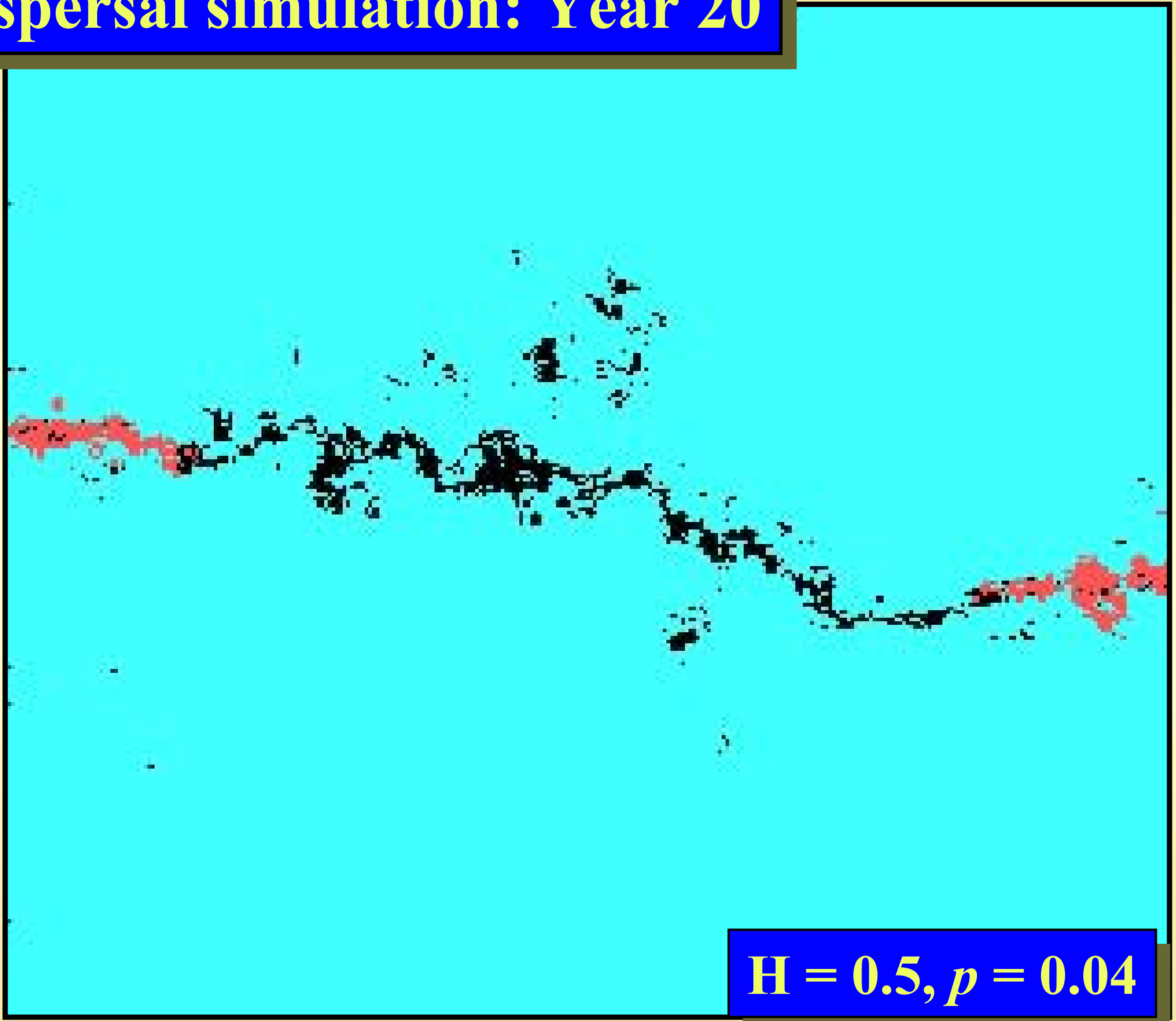
H = 0.5, p = 0.04

Dispersal simulation: Year 10



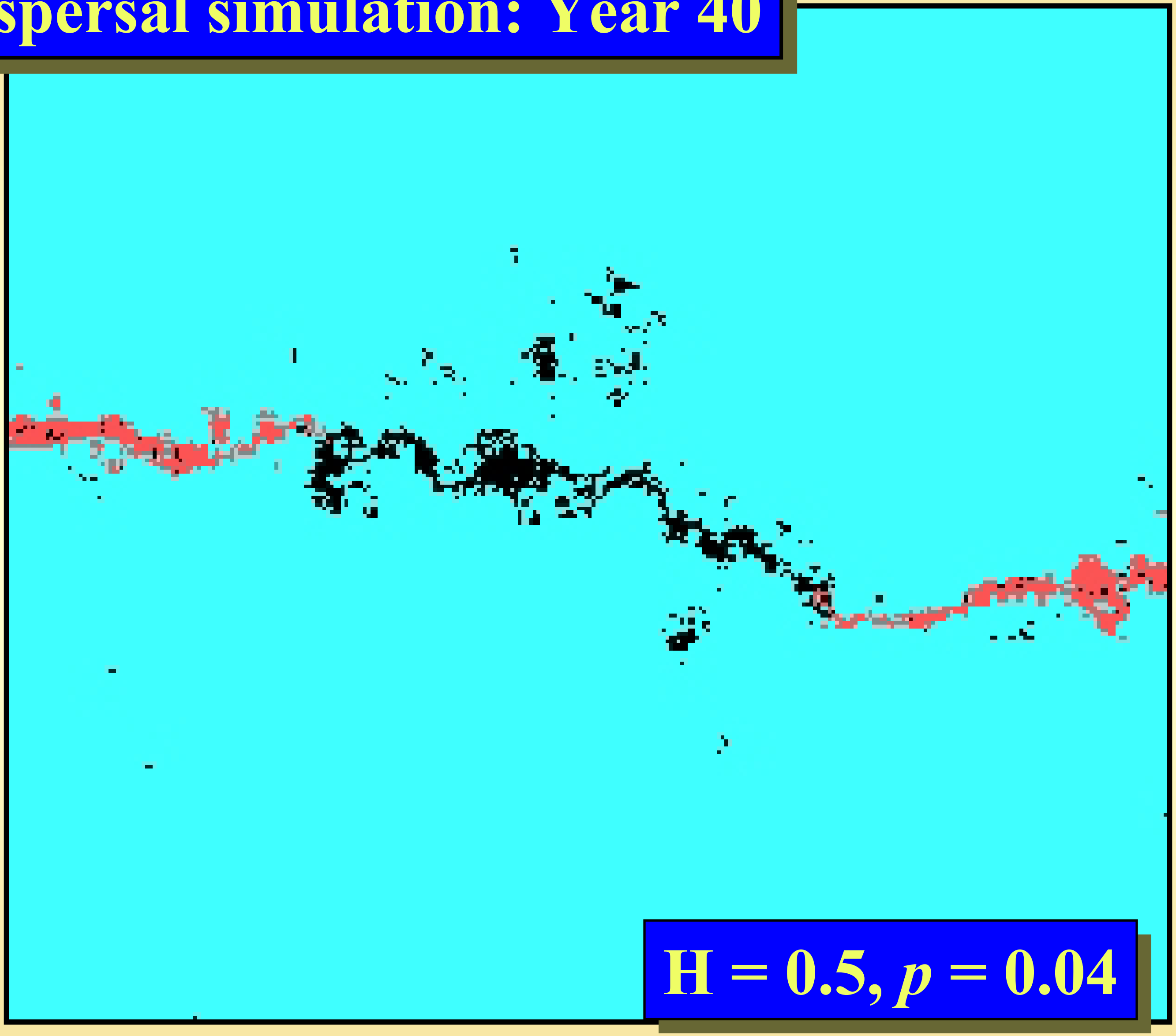
H = 0.5, p = 0.04

Dispersal simulation: Year 20



H = 0.5, p = 0.04

Dispersal simulation: Year 40



H = 0.5, p = 0.04

Factorial experiment (N = 810 simulations)

- ***H and p***
  - *Control landscape pattern*
- ***Fecundity***
  - *Rate of invasion is known to be dependent on fecundity*
- ***Niche width***
  - *Invading species not restricted to corridor*
- ***Competition***
  - *Residents species provides resistance*

ANOVA (expressed as % effect)

Factor	*df*	*v*	*S*
H	2	42.3	41.6
p	1	16.3	10.1
R	1	14.8	1.0
Niche	1	4.3	13.4
Comp	1	5.6	1.6
R^2		84.3	73.7

N = 810 simulations

Landscape effect

- *Explained 58% of the variance in invasion speed*
  - *51% of variance in pattern of occupied habitat*
- parameters *H* and *p* sufficient to “explain” landscape effects
  - other landscape indices not more predictive
  - …although other indices may be easier to estimate
- thus, a few measure of *landscape structure* may be sufficient to predict conditions for invasion

Species parameters

- fecundity more important than niche width
  - for ***v***, the rate invasion
  - but not for ***S***, the pattern of habitat occupied
- competition had limited effect
  - probably because the central portion of the corridor was always unoccupied

- further testing required …. but

Is this surprising?

- ***Simple relationships*** between pattern combined with initial estimates of invasion speed may be adequate to predict invasion success
- ***Obviously****:* Empty, continuous corridors provide optimal invasion routes
  - “empty” and “continuous” are relative to the scale of invader’s dispersal kernal
  - expect increases in invasion following disturbances which eliminate competitors

Some recommendations (1)

- All landscapes should be described (metadata) with a minimum sets of metrics and statistics, including:
  - grain and extent
  - data source, land-cover classification rule, caveats for wise use
  - number (N) and amount (p_i) of each class type

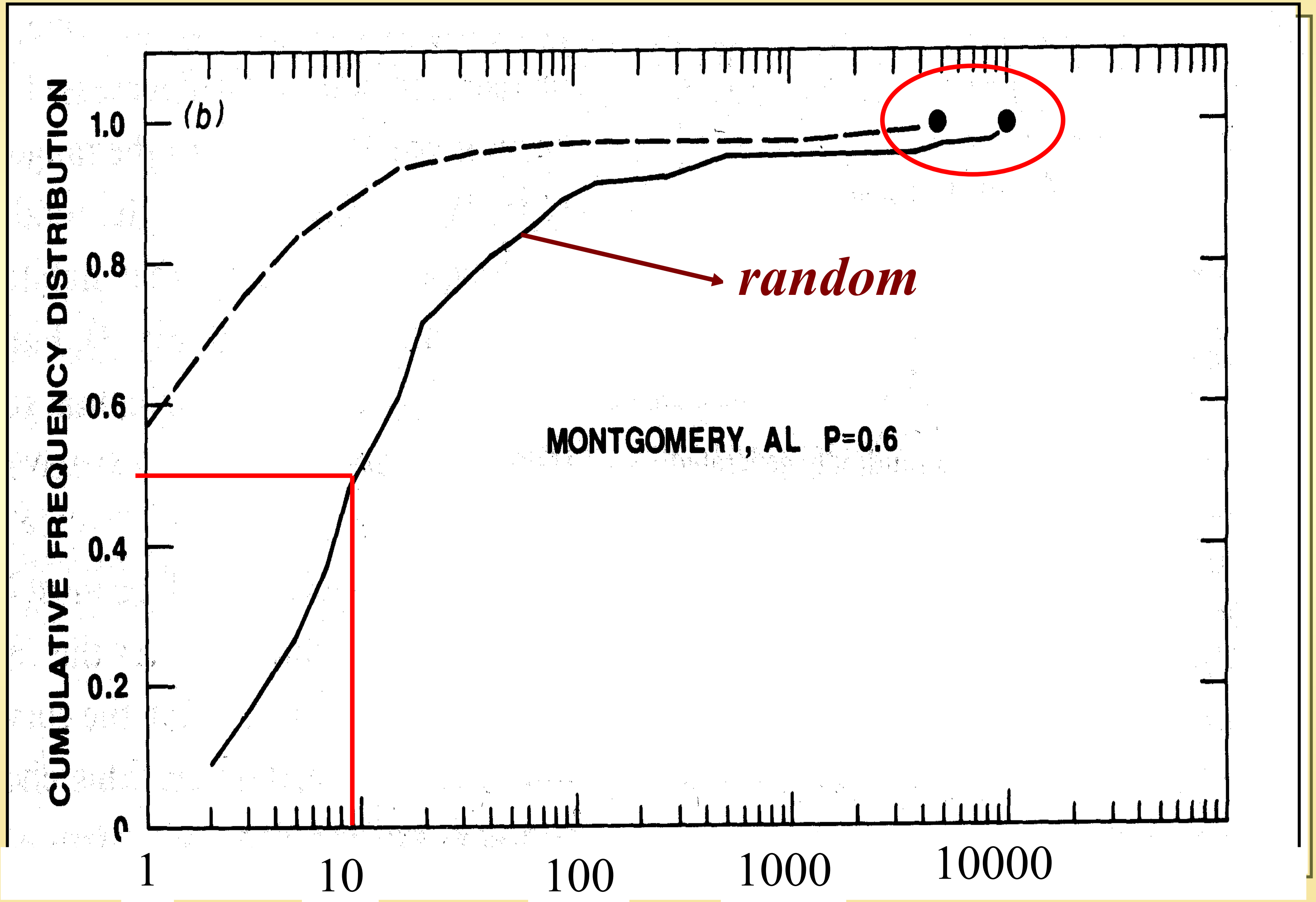
Some recommendations (2)

- Metrics should be consistently evaluated
  1. Are they sensitive (and monotonic) to pattern change?
  2. How are they affected by scale (grain and extent of map)?
  3. What are the confidence limits (i.e., how much change is needed to be significant)?
  4. Clear, mathematical description with analytic comparison to related metrics <u>and methods</u> should be available
  - Computer code for calculations made available

Some recommendations (3)

- Innovations required.
  - Why not rely more on established statistical methods to evaluate patterns?
    - i.e., test differences in distributions of patch sizes. These distributions are non-normal and simple averages do not work.
    - Monte Carlo methods (such as bootstrapping) are under-utilized

Comparison of random with actual map



And then there are truth tables!

0 = absent
1 = present

Observed \ Predicted	0	1
0	**true negative**	false positive
1	false negative	**true positive**

Truth table: power analysis

0 = absent
1 = present

		Predicted 0	Predicted 1
Observed	0	$1-\alpha$	α
	1	β	$1-\beta$